IBM SP System

IBM SP

- SP system simultaneously brings dozens to hundreds of RISC processors nodes to a computing problem
- Nodes
  - PowerPC 604e : SMP 332 Mhz
    - Thin : 2 or 4 way
    - Wide : 2 or 4 way
    - High : 2, 4, or 8 way
  - P2SC
    - Thin 120 Mhz
    - Wide 135 Mhz
    - Thin 160 Mhz (SDSC nodes)
- All nodes connected by high performance switch

IBM SP System (cont.)

- PowerPC 604e SMP nodes have both L1 and L2 cache and better I/O rate to disk
  - Thin and Wide nodes
    - L1 cache: 32 KB data / 32 KB instr
    - L2 cache : 256 KB
  - High nodes
    - L1 : 32 KB data / 32 KB instr
    - L2 : 2 MB
- P2SC nodes have only L1 cache (128 KB data / 32 KB instr) and better FP performance than PowerPC

P2SC Processor

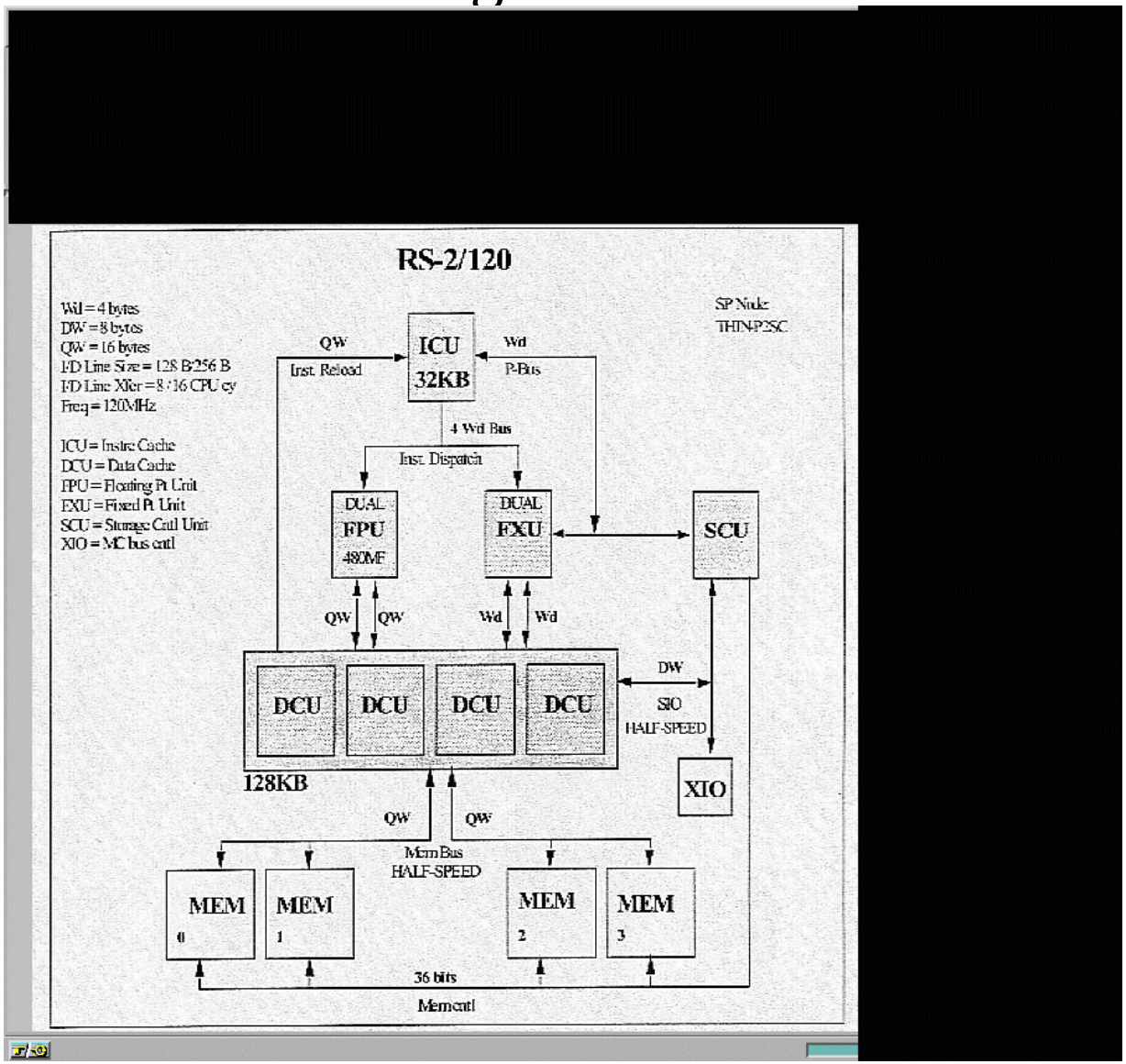
(P2SC nodes currently at SDSC; and will be upgraded to PowerPC SMP nodes in mid 1999)

- Power2 Super Chip (P2SC) thin node with 160 Mhz
- Superscalar - issues upto 6 (branch, conditional, 2 FP, and 2 INT) instructions/CP
- Peak performance 640 MFLOPS
- On chip 128 Kbytes and 4-way set associative data cache
- On chip 32KB and 2-way set associative ICACHE

Floating Point Execution Unit (P2SC)

- The FPU contains:
  - 32 64-bit IEEE floating point registers
  - FPU with dual floating point execution units
  - Has the capability to execute a floating point multiply-add (FMA) instruction with a latency of two cycles and to start independent FMA every cycle
  - Special instruction for increased performance:
    - Floating point square root instruction performed in hardware
    - Quad-word floating point storage reference allows two adjacent floating point registers to be loaded with two adjacent floating point numbers with one instruction

CPU Block diagram of IBM SP P2SC

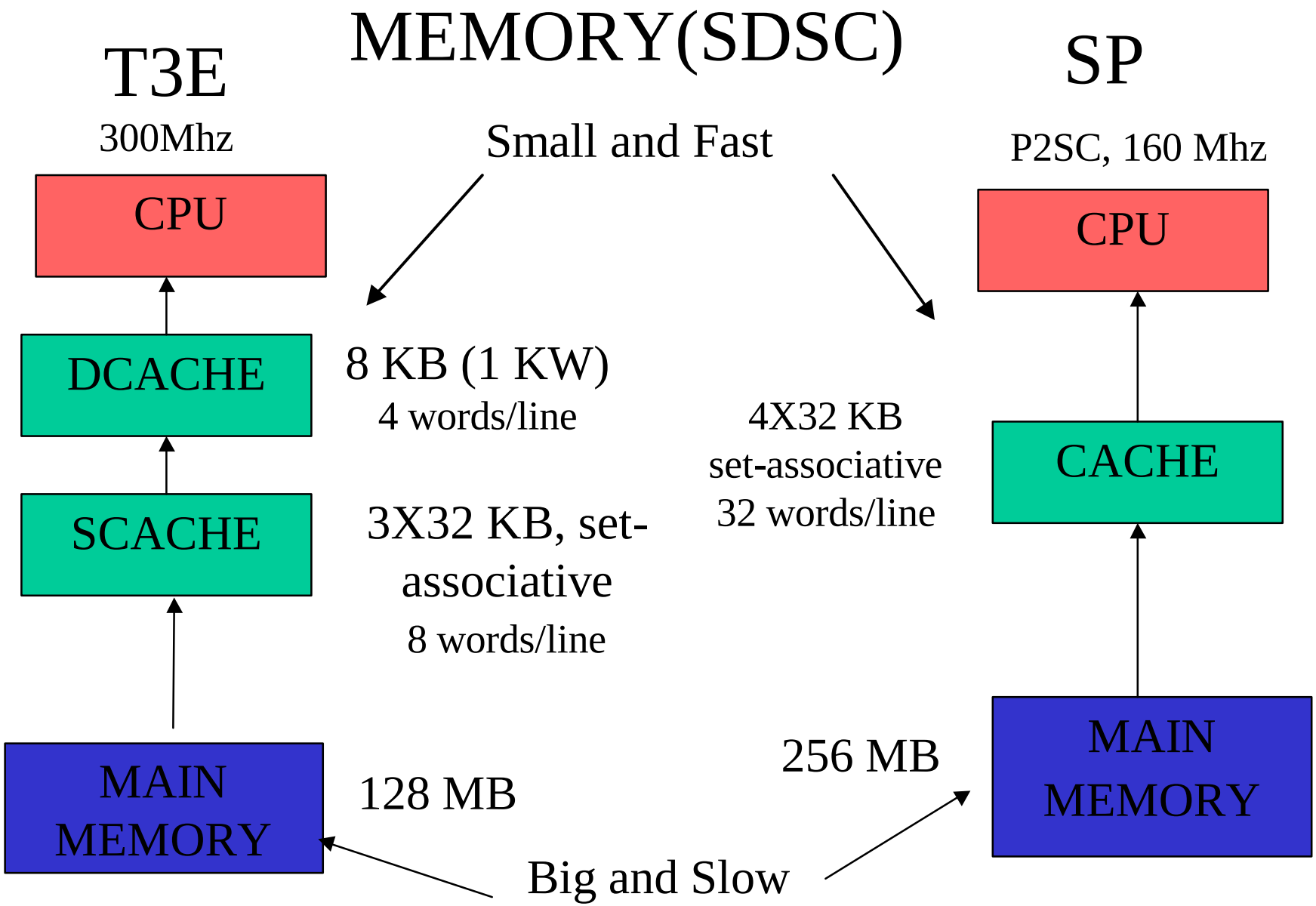


DCACHE (P2SC)

- On-chip data cache (only L1 cache as opposed to L1 and L2 cache on the Cray T3E)
- Cache size is 128 Kbytes
- 4 way set associative, 4 cache sets of 32 Kbytes each
- 256 Bytes per line or 32 words per line

ICACHE (P2SC)

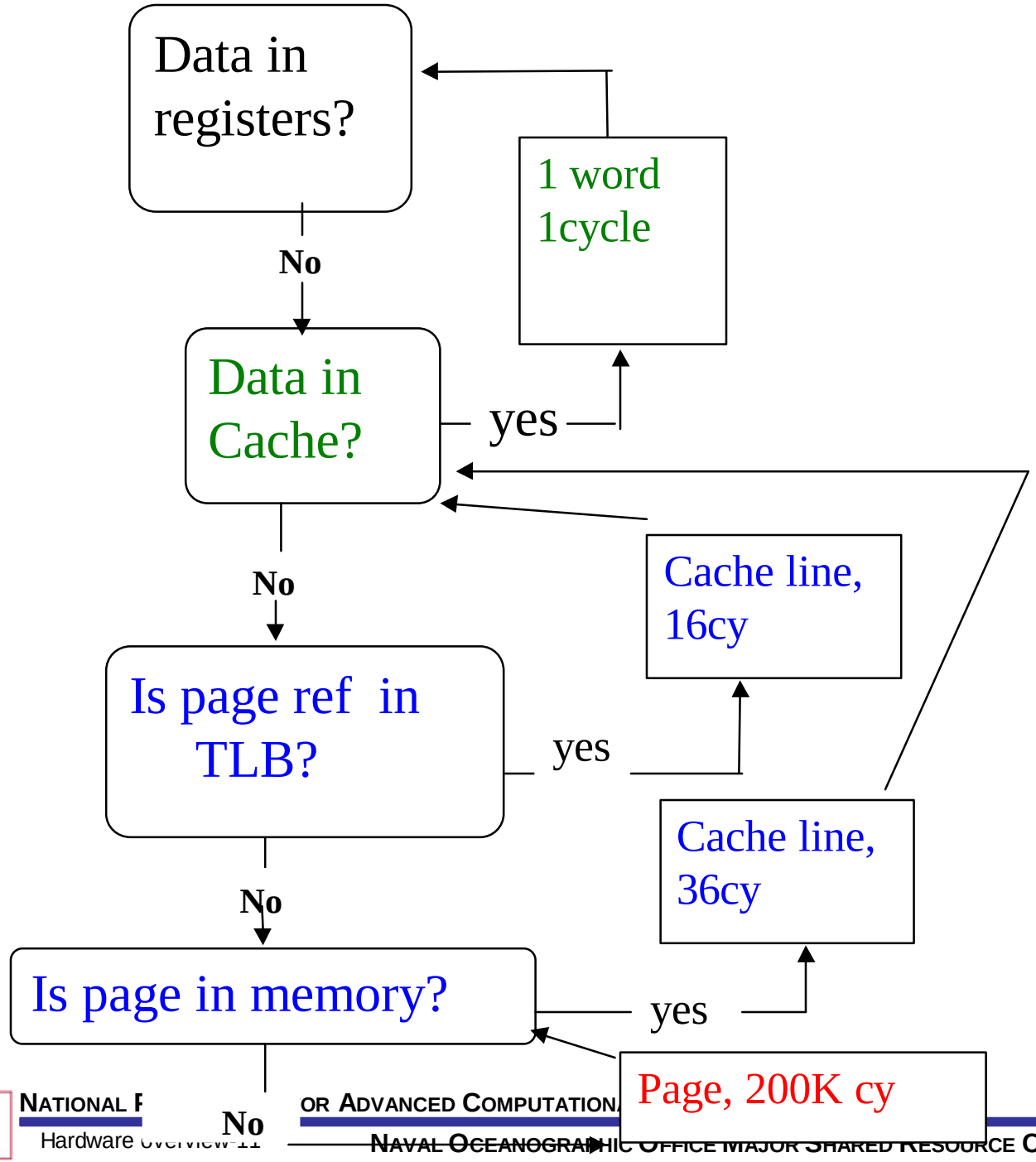
- 32 Kbytes
- 2-way set associative


MEMORY(SDSC)
T3E
300Mhz
Small and Fast
SP
P2SC, 160 Mhz
CPU
CPU
DCACHE
8 KB (1 KW)
4 words/line
4X32 KB
set-associative
32 words/line
CACHE
SCACHE
3X32 KB, set-
associative
8 words/line
MAIN
MEMORY
128 MB
256 MB
MAIN
MEMORY
Big and Slow

Memory Characteristics of IBM SP

- SP is a virtual memory machine and allows data structure to be larger than main memory's capacity
- A page (4096 bytes) in main memory that is recently been accessed may have its address listed in TLB which allows faster access
- Swap space is the area in disk that is set aside for the portions of virtual memory that won't fit in the real memory

- Next slide provides memory latency diagram for IBM SP P2SC 120 Mhz


Data in registers?
1 word 1cycle
No
Data in Cache?
yes
No
Cache line, 16cy
Is page ref in TLB?
yes
No
Cache line, 36cy
Is page in memory?
yes
Page, 200K cy
No

SP High Performance Switch

- Provides the internal message passing fabric that connects all of the SP processors together
- SP networks are bidirectional multistage interconnection networks
- Bidirectional any-to-any internode connection allows all nodes to send message simultaneously
- Supports multi-user environment - multiple jobs may run simultaneoulsy over the switch (one user does not monopolize switch)

SP High Performance Switch (cont..)

- Switch has two protocols either of which can be requested by user:
  - Internet Protocol : default; permits shared usage of HPS by multiple processes
  - User Space (US) : intended for parallel applications that require optimum communication performance. Only one process per node may use US communication.

Latency and Bandwidth on SP Network

Switch type	**Latency (micro sec)**	**Bandwidth (MB/sec)**
IP	**~ 300**	**~ 25**
US	**~ 40**	**100**

IBM's RS6000 Web Page

- http://www.rs6000.ibm.com/sp.html